

Viewpoint

# Potential Roles of Large Language Models in the Production of Systematic Reviews and Meta-Analyses

Xufei Luo<sup>1,2,3,4,5</sup>; Fengxian Chen<sup>6</sup>, PhD; Di Zhu<sup>7</sup>, MPH; Ling Wang<sup>7</sup>, MPH; Zijun Wang<sup>1,2,3,4,5</sup>; Hui Liu<sup>1,2,3,4,5</sup>; Meng Lyu<sup>7</sup>, MPH; Ye Wang<sup>7</sup>, MPH; Qi Wang<sup>8,9</sup>, PhD; Yaolong Chen<sup>1,2,3,4,5</sup>, MD, PhD

<sup>1</sup>Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

<sup>2</sup>World Health Organization Collaboration Center for Guideline Implementation and Knowledge Translation, Lanzhou, China

<sup>3</sup>Institute of Health Data Science, Lanzhou University, Lanzhou, China

<sup>4</sup>Key Laboratory of Evidence Based Medicine and Knowledge Translation of Gansu Province, Lanzhou University, Lanzhou, China

<sup>5</sup>Research Unit of Evidence-Based Evaluation and Guidelines, Chinese Academy of Medical Sciences (2021RU017), School of Basic Medical Sciences, Lanzhou University, Lanzhou, China

<sup>6</sup>School of Information Science & Engineering, Lanzhou University, Lanzhou, China

<sup>7</sup>School of Public Health, Lanzhou University, Lanzhou, China

<sup>8</sup>Department of Health Research Methods, Evidence and Impact, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

<sup>9</sup>McMaster Health Forum, McMaster University, Hamilton, ON, Canada

**Corresponding Author:**

Yaolong Chen, MD, PhD

Evidence-Based Medicine Center

School of Basic Medical Sciences

Lanzhou University

No 199 Donggang West Road

Chengguan District

Lanzhou, 730000

China

Phone: 86 13893104140

Email: [chevidence@lzu.edu.cn](mailto:chevidence@lzu.edu.cn)

## Abstract

Large language models (LLMs) such as ChatGPT have become widely applied in the field of medical research. In the process of conducting systematic reviews, similar tools can be used to expedite various steps, including defining clinical questions, performing the literature search, document screening, information extraction, and language refinement, thereby conserving resources and enhancing efficiency. However, when using LLMs, attention should be paid to transparent reporting, distinguishing between genuine and false content, and avoiding academic misconduct. In this viewpoint, we highlight the potential roles of LLMs in the creation of systematic reviews and meta-analyses, elucidating their advantages, limitations, and future research directions, aiming to provide insights and guidance for authors planning systematic reviews and meta-analyses.

(*J Med Internet Res* 2024;26:e56780) doi: [10.2196/56780](https://doi.org/10.2196/56780)

**KEYWORDS**

large language model; ChatGPT; systematic review; chatbot; meta-analysis

## Introduction

A systematic review is the result of a systematic and rigorous evaluation of evidence, which may or may not include a meta-analysis [1]. Owing to the strict methodology and comprehensive summary of evidence, high-quality systematic reviews are considered the highest level of evidence, positioned at the top of the evidence pyramid [2]. Additionally, high-quality

systematic reviews and meta-analyses are often used to support the development of clinical practice guidelines, aid clinical decision-making, and inform health care policy formulation [3]. Currently, the methods of systematic reviews and meta-analyses are applied in various disciplines in medicine and beyond such as law [4], management [5], and economics [6], and have yielded positive results, contributing to the continuous advancement of these fields [7].

The process of conducting systematic reviews demands a substantial investment in terms of time, resources, human effort, and financial capital [8]. To expedite the development of systematic reviews and meta-analyses, various automated or semiautomated tools such as Covidence have been developed [9,10]. However, the emergence of large language models (LLMs), particularly chatbots such as GPT, presents a set of both challenges and opportunities in the realm of systematic reviews and meta-analyses [11]. Based on the emerging literature in this field, we here provide our perspectives on the potential for harnessing the capabilities of LLMs to accelerate the production of systematic reviews and meta-analyses, while also scrutinizing the potential impacts and delineating the crucial steps involved in this process.

### The Process and Challenges of Performing a Systematic Review and Meta-Analysis

The procedures and workflows for conducting systematic reviews and meta-analyses are well-established. Currently, researchers often refer to the Cochrane Handbooks recommended by the Cochrane Library for intervention or diagnostic reviews [12,13]. In addition, some scholars and institutions have developed detailed guidelines on the steps and methodology for performing systematic reviews and meta-analyses [14-17]. Generally speaking, researchers should take the following steps to produce a high-quality systematic review and meta-analysis: determine the clinical question,

register and draft a protocol, set inclusion and exclusion criteria, develop and implement a search strategy, screen the literature, extract data from included studies, assess the quality and risk of bias of included studies, analyze and processed data, write up the full text of the manuscript, and submit the manuscript for publication, as illustrated in Figure 1. These different steps contain many subtasks; therefore, conducting a complete systematic review and meta-analysis requires fairly complex and time-consuming work.

Although systematic reviews and meta-analyses have been widely applied and play an important role in developing guidelines and informing clinical decision-making, their production process faces many challenges. One of these challenges is the long production time and large resource requirements. The average estimated time to complete and publish a systematic review is 67.3 weeks, requiring 5 researchers and costing approximately US \$140,000 [18,19]. More recently, the development of automated and semiautomated tools using natural language processing and machine learning have accelerated systematic review and meta-analysis production to some extent [20], with studies showing that such tools can help to produce a systematic review and meta-analysis within 2 weeks [21]. However, these tools also have some limitations. First, no single tool can fully accelerate the entire production process of systematic reviews and meta-analyses. Second, these tools cannot process and analyze literature written in different languages. Finally, the reliability of the results generated by these automated and semiautomated tools needs further validation as they are not yet widely adopted for this purpose.

Figure 1. The process of conducting a systematic review and meta-analysis.



## Applications of LLMs in Medical Research

Chatbots based on LLMs such as ChatGPT, Google Gemini, and Claude have become widely applied in medical research. These chatbots have proven to be valuable in tasks ranging from knowledge retrieval, language refinement, content generation, and medical exam preparation to literature assessment. ChatGPT has been shown to excel in accuracy, completeness, nuance, and speed when generating responses to clinical inquiries in psychiatry [22]. Moreover, LLMs such as ChatGPT play a pivotal role in automating the evaluation of medical literature, facilitating the identification of accurately reported research findings [23]. Despite their significant contributions, these chatbots are not without limitations. Challenges such as the

potential for generating misleading content and susceptibility to academic deception necessitate further scholarly discourse on effective mitigation strategies. Standardized reporting practices may contribute to delineating the applications of ChatGPT and mitigating research biases [24].

ChatGPT has also demonstrated significant application potential and promise in the process of conducting systematic reviews and meta-analyses. Various studies [11,25-32] indicate that ChatGPT can play a pivotal role in formulating clinical questions, determining inclusion and exclusion criteria, screening literature, assessing publications, generating meta-analysis code, and assisting the full-text composition, among other relevant tasks. The details of these capabilities are summarized in [Table 1](#).

**Table 1.** The possible functions of chatbots in the creation of systematic reviews and meta-analyses encompassing separate stages of the process.

Tasks	Potential roles and application steps of chatbots	References
Determine the research topic/question	<ul style="list-style-type: none"> <li>Identify previously published systematic reviews and meta-analyses on the same topic.</li> <li>Assist in determining the rationale for the research question.</li> <li>Clarify the PICO (Population, Intervention, Comparison, Outcome) question.</li> </ul>	[11,33-35]
Register and write a research proposal	<ul style="list-style-type: none"> <li>Generate preliminary, unverified registration information.</li> <li>Draft an initial research proposal, subject to validation.</li> </ul>	[11,36,37]
Define inclusion and exclusion criteria	<ul style="list-style-type: none"> <li>Establish inclusion criteria.</li> <li>Establish exclusion criteria.</li> </ul>	[11,38]
Develop a search strategy and conduct searches	<ul style="list-style-type: none"> <li>Develop and optimize search strategies.</li> <li>Implement retrieval.</li> <li>Collect grey literature.</li> </ul>	[11,25,29,33,39-42]
Screen the literature	<ul style="list-style-type: none"> <li>Remove duplicate records.</li> <li>Screen literature titles, abstracts, and keywords.</li> <li>Screen the full text of the obtained literature.</li> <li>Download the full text of the literature.</li> </ul>	[11,25,27,28,33,34,43-47]
Extract the data	<ul style="list-style-type: none"> <li>Extract basic information.</li> <li>Extract patient information.</li> <li>Extract outcome information.</li> <li>Extract table information.</li> </ul>	[11,25,26,47-50]
Assess the risk of bias	<ul style="list-style-type: none"> <li>Extract relevant information based on the scale.</li> <li>Evaluate the risk of bias based on the scale.</li> <li>Present visual results.</li> </ul>	[26,51-53]
Analyze the data/meta-analyses	<ul style="list-style-type: none"> <li>Extract outcome information.</li> <li>Generate figures and tables for some results.</li> </ul>	[11,26,37,54]
Draft the full manuscript	<ul style="list-style-type: none"> <li>Search for relevant references.</li> <li>Polish language and grammar.</li> <li>Adjust the reference citation format.</li> <li>Summarize the abstract.</li> </ul>	[25,33,55-58]
Submit and publish	<ul style="list-style-type: none"> <li>Assist in selecting a suitable journal.</li> <li>Adjust the manuscript format.</li> <li>Compose a cover letter.</li> <li>Assist in preparing the submission.</li> </ul>	[33,59]

## ***Potential Roles of LLMs in Producing Systematic Reviews and Meta-Analyses***

### **Determine the Research Topic/Question**

Determining the clinical question of interest represents the initial and paramount step in the process of conducting systematic reviews and meta-analyses. At this juncture, it is crucial to ascertain whether comparable systematic reviews and meta-analyses have already been published and to delineate the scope of the forthcoming review and meta-analysis. Generally, for interventional systematic reviews, the Patient, Intervention, Comparison, and Outcome (PICO) framework is considered for defining the scope and objectives of the research question [60]. In this context, ChatGPT serves a dual role. On the one hand, it expeditiously aids in searching for published systematic reviews and meta-analyses related to the relevant topics (see [Multimedia Appendix 1](#) and [Multimedia Appendix 2](#)) [34]. On the other hand, ChatGPT assists in refining the clinical question that needs to be addressed (see [Multimedia Appendix 3](#)), facilitating prompt determination of the feasibility of undertaking the proposed study. However, it is important to be cautious of the retrieval of false literature [35].

### **Register and Write a Research Proposal**

The registration and proposal writing process constitutes a pivotal preparatory phase for conducting systematic reviews and meta-analyses. Registration enhances research transparency, fosters collaboration among investigators, and mitigates the redundancy of research endeavors. Drafting a proposal helps in elucidating the research objectives and methods, providing robust support for the smooth execution of the study. For LLMs, generating preliminary registration information and initial proposal content is remarkably convenient and facile (see [Multimedia Appendix 4](#) and [Multimedia Appendix 5](#)). For example, ChatGPT can assist researchers in generating the statistical methods for a research proposal [37]. However, considering that LLMs often generate fictitious literature, the content they produce may be inaccurate; thus, discernment and validation of the generated content remain essential considerations.

### **Define Inclusion and Exclusion Criteria**

The inclusion and exclusion criteria for systematic reviews and meta-analyses are instrumental in determining the screening standards for studies. Therefore, strict and detailed inclusion and exclusion criteria contribute to the smooth and high-quality conduct of preparing systematic reviews and meta-analyses. The use of a chatbot based on LLMs can help in establishing the inclusion and exclusion criteria (see [Multimedia Appendix 6](#)) [38]; however, the inclusion criteria need to be optimized and adjusted according to the specific research objectives and the exclusion criteria should be based on the foundation of the inclusion criteria. Therefore, manual adjustments and optimizations are also necessary.

### **Develop a Search Strategy and Conduct Searches**

ChatGPT can assist in formulating search strategies, using PubMed as an example [40]. Researchers can simply list their

questions using the PICO framework and a search strategy can be quickly generated ([Multimedia Appendix 1](#) and [Multimedia Appendix 2](#)). Based on the generated search strategy, one method is to copy the strategy from ChatGPT and paste it into the PubMed search box for direct retrieval [40,41]. Another approach involves using the OpenAI application programming interfaces (APIs) to invoke PubMed APIs with the search strategy generated by ChatGPT. This facilitates searching the PubMed database, obtaining search results, and applying predetermined inclusion and exclusion criteria. Subsequently, ChatGPT can be used to filter the search results, exporting and recording the filtered results in JSON format. This integrated process encompasses search strategy formulation, retrieval, and filtering. However, the direct use of LLMs to generate search strategies and complete the one-stop process of searching and screening may not yet be mature, and this poses a significant challenge for generating the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flowchart. Therefore, we suggest using LLMs to generate search strategies, which should then be optimized and modified by librarians and computer experts (specializing in LLMs) before manually searching the databases. Additionally, to use search strategies transparently and reproducibly, the detailed prompts used should be reported [40,42].

### **Screen the Literature**

Literature screening is one of the most time-consuming steps in the creation of systematic reviews and meta-analyses. Prior to the advent of ChatGPT, there were already many automated and semiautomated tools available for literature screening, such as Covidence, EPPI-Reviewer, DistillerSR, and others [39]. With the emergence of ChatGPT, researchers can now train the model based on predefined inclusion criteria. Subsequently, ChatGPT can be used to automatically screen records retrieved from databases and obtain the filtered results. Previous studies suggested that using ChatGPT in the literature selection process for a meta-analysis substantially diminishes the workload while preserving a recall rate on par with that of manual curation [28,44-47].

### **Extract the Data**

Data extraction involves obtaining information from primary studies and serves as a primary source for systematic reviews and meta-analyses. Generally, when conducting systematic reviews and meta-analyses, basic information must be extracted from the original studies, such as publication date, country of conduct, and the journal of publication. Additionally, characteristics of the population, such as patient samples, age, gender/sex, and outcome data, are also extracted, including event occurrences, mean change values, and total sample size. Currently, tools based on natural language processing and LLMs, such as ChatGPT and Claude, demonstrate high accuracy in extracting information from PDF documents (see [Multimedia Appendix 7](#) for an example) [47-50]. However, it is important to note that despite their promising capabilities, manual verification remains a necessary step in the data extraction process when using these artificial intelligence (AI) tools [61]. Using LLMs to extract data can help avoid random errors;



however, caution is still required when extracting data from figures or tables [47-50].

### Assess the Risk of Bias

Assessing the bias of risk involves evaluating the internal validity of studies included in research. For randomized controlled trials, tools such as Risk of Bias (RoB) [62] or its updated version RoB 2 [63] are typically used, with an estimated review time of 10-15 minutes per trial. However, automated tools such as RobotReviewer can streamline the extraction and evaluation process in batches [51-53], thereby improving efficiency, although manual verification is still necessary. Additionally, chatbots based on LLMs can aid in risk of bias assessment (see [Multimedia Appendix 8](#)), and their accuracy appears to be comparable to that of human evaluations [23].

### Analyze the Data/Meta-Analysis

Data analysis serves as the source of systematic review results, typically encompassing basic information and outcome findings. The meta-analysis may be one outcome, along with potential components such as subgroup analysis, sensitivity analysis, meta-regression, and detection of publication bias. Numerous software options are available to facilitate these data analyses, including Stata, RevMan, Rstudio, and others [43]. Currently, it appears that chatbots based on LLMs may not fully execute data analysis independently, although they can extract the relevant information. Subsequently, one can employ corresponding software for comprehensive data analysis. Alternatively, after extracting information with chatbots, the ChatGPT Code Interpreter can assist in analysis and generating graphical results, although this requires a subscription to ChatGPT Plus. Moreover, an LLM markedly accelerates the data analysis process, empowering researchers to handle larger data sets with greater efficacy [54].

### Draft the Full Manuscript

The complete drafting of systematic reviews and meta-analyses should adhere to the PRISMA reporting guidelines [64]. It is not advisable to use chatbots such as ChatGPT for article composition. On the one hand, the accuracy and integrity of content generated by ChatGPT require human verification. On the other hand, various research types and journals have different requirements for full-length articles, making it challenging to achieve uniformity in the generated content. However, using tools such as GPT for language refinement and adjusting the content logic can be considered to enhance the quality and readability of the article [33,55]. It is important to declare the use of GPT-related tools in the methods, acknowledgments, or appendices of the article to ensure transparency [24,65].

### Submit and Publish

Submission and publication represent the final steps in the process of conducting systematic reviews and meta-analyses, aside from subsequent updates. At this stage, the potential role of LLM-based tools is to assist authors in recommending suitable journals (see [Multimedia Appendix 9](#)). These tools might also aid in crafting components required along with submission of the manuscript such as cover letters and highlights [59]. However, it is imperative to emphasize that the content generated by these tools requires manual verification to ensure

accuracy, and all authors should be accountable for the content generated by LLMs.

## Benefits and Drawbacks of Using LLMs

Systematic reviews and meta-analyses are crucial evidence types that support the development of guidelines [3]. The benefits of employing LLM-based chatbots in the production of systematic reviews and meta-analyses include increased speed, such as in the stages of evidence searching, data extraction, and assessment of bias risk; these tools can also enhance accuracy by reducing human errors such as those made while extracting essential information and pooling data. However, there are also drawbacks of these applications of LLMs, such as the potential for generating hallucinations, the requirement for human verification owing to the poor reliability of the models, and that the entire systematic review process is not replicable. Moreover, when interacting with LLM chatbots, it is important to manage data privacy. In particular, when using LLMs to analyze data, especially when including personal patient information, ethical approval and management must be properly addressed.

## Challenges and Solutions

While LLMs can assist in accelerating the production of systematic reviews and meta-analyses in some steps, enhancing accuracy and transparency, and saving resources, they also face several challenges. For instance, LLMs cannot promptly update their versions and information. For example, ChatGPT 3.5 has been trained on data available in 2021. Thus, limitations such as the length of prompts and token constraints, as well as restrictions related to context associations, may potentially impact the overall results and user experience [25]. Although LLM-based autonomous agents have made strides in tasks related to systematic reviews and meta-analyses, their applications are still associated with various issues related to personalization, updating knowledge, strategic planning, and complex problem-solving. The development of LLM-driven autonomous agents adept at systematic reviews and meta-analyses warrants further exploration [66]. The use of LLMs as centrally controlled intelligent agents encompasses the ability to handle precise literature screening, extract and analyze complex data, and assist in manuscript composition, as highlighted by proof-of-concept demonstrations such as MetaGPT [67]. Moreover, the continuous growth of the use of LLMs can pose a significant challenge in ensuring the accuracy of information provided in systematic reviews, particularly if LLMs are indiscriminately overused.

To better facilitate the use of tools such as ChatGPT in systematic reviews and meta-analyses, we believe that, first and foremost, authors should understand the scope and scenarios for applying ChatGPT, clearly defining which steps can benefit from these tools. Second, for researchers, collaboration with computer scientists and AI engineers is crucial to optimize the prompts and develop integrated tools based on LLMs, such as web applications. These tools can assist in seamless transitions between different tasks in the systematic review process. Lastly, for journal editors, collaboration with authors and reviewers is essential to adhere to reporting and ethical principles associated

with the use of GPT and similar tools [24,68]. This collaboration aims to promote transparency and integrity, while preventing indiscriminate overuse in the application of LLMs in systematic reviews and meta-analyses.

## ***Future Perspectives and Conclusion***

The emergence of LLMs could have a significant impact on the production of systematic reviews and meta-analyses. In this process, the application of chatbots such as ChatGPT has the potential to speed up certain steps such as literature screening, data extraction, and risk of bias assessment, which are processes that typically consume a considerable amount of time. However,

it is important to note that if AI methods such as GPT are employed in performing systematic reviews, disclosure and declaration of the use of these tools are essential. This includes specifying the AI tools used, their roles, and the areas of application within the review process, among other relevant information for full disclosure [24]. In this context, developing a reporting guideline is warranted to guide the application of LLM tools in systematic reviews and meta-analyses. Although the PRISMA 2020 guideline briefly addresses the use of automation technologies, its coverage is limited to steps such as screening, and there is a lack of comprehensive guidance on the broader spectrum of applications [64].

## **Acknowledgments**

ChatGPT 3.5 designed by OpenAI was used to help with language editing. The authors take the ultimate responsibility for the content of this publication.

## **Authors' Contributions**

XL and YC were responsible for conceptualization of the article. XL, FC, DZ, and LW generated the examples with the large language models and wrote the first draft of the article. XL, ZW, HL, ML, YW, QW, and YC reviewed and edited the manuscript. YC supervised the study, takes full responsibility for the work and conduct of the study, has access to the data, and controlled the decision to publish. All authors read the final manuscript and approved the submission.

## **Conflicts of Interest**

None declared.

## **Multimedia Appendix 1**

Using ChatGPT 4.0 to assist in generating PubMed search strategies for assessing systematic reviews.  
[\[PNG File , 3022 KB-Multimedia Appendix 1\]](#)

## **Multimedia Appendix 2**

The results obtained after searching the PubMed database based on the search strategy generated by ChatGPT.  
[\[PNG File , 1409 KB-Multimedia Appendix 2\]](#)

## **Multimedia Appendix 3**

Using ChatGPT 4.0 to assist in optimizing the clinical question for conducting a systematic review and meta-analysis.  
[\[PNG File , 210 KB-Multimedia Appendix 3\]](#)

## **Multimedia Appendix 4**

Using ChatGPT 4 to generate PROSPERO (International Prospective Register of Systematic Reviews) registration information.  
[\[PNG File , 482 KB-Multimedia Appendix 4\]](#)

## **Multimedia Appendix 5**

Proposal of a systematic review and meta-analysis related to exercises for osteoarthritis generated by Claude 3 based on the provided prompts.  
[\[PNG File , 1435 KB-Multimedia Appendix 5\]](#)

## **Multimedia Appendix 6**

The inclusion and exclusion criteria for a systematic review and meta-analysis on exercise therapy for osteoarthritis based on GPT-4.  
[\[PNG File , 2225 KB-Multimedia Appendix 6\]](#)

## Multimedia Appendix 7

Using Claude 3 for data extraction from PDF documents: an example with three randomized controlled trials.

[\[PNG File , 3184 KB-Multimedia Appendix 7\]](#)

## Multimedia Appendix 8

Using Claude 3 for risk of bias assessment: an example with two randomized controlled trials.

[\[PNG File , 4317 KB-Multimedia Appendix 8\]](#)

## Multimedia Appendix 9

Using GPT-4 to assist in selecting target journals for submission of a systematic review and meta-analysis.

[\[PNG File , 203 KB-Multimedia Appendix 9\]](#)

## References

1. Jahan N, Naveed S, Zeshan M, Tahir MA. How to conduct a systematic review: a narrative literature review. *Cureus*. Nov 04, 2016;8(11):e864. [[FREE Full text](#)] [doi: [10.7759/cureus.864](https://doi.org/10.7759/cureus.864)] [Medline: [27924252](https://pubmed.ncbi.nlm.nih.gov/27924252/)]
2. Wallace SS, Barak G, Truong G, Parker MW. Hierarchy of evidence within the medical literature. *Hosp Pediatr*. Aug 01, 2022;12(8):745-750. [doi: [10.1542/hpeds.2022-006690](https://doi.org/10.1542/hpeds.2022-006690)] [Medline: [35909178](https://pubmed.ncbi.nlm.nih.gov/35909178/)]
3. Institute of Medicine, Board on Health Care Services, Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, Graham R, Mancher M, Miller Wolman D, et al. *Clinical Practice Guidelines We Can Trust*. Washington, DC: National Academies Press; 2011.
4. Bystranowski P, Janik B, Próchnicki M, Skórska P. Anchoring effect in legal decision-making: a meta-analysis. *Law Hum Behav*. Feb 2021;45(1):1-23. [doi: [10.1037/lhb0000438](https://doi.org/10.1037/lhb0000438)] [Medline: [33734746](https://pubmed.ncbi.nlm.nih.gov/33734746/)]
5. Geyskens I, Krishnan R, Steenkamp JEM, Cunha PV. A review and evaluation of meta-analysis practices in management research. *J Management*. Feb 05, 2008;35(2):393-419. [doi: [10.1177/0149206308328501](https://doi.org/10.1177/0149206308328501)]
6. Bagepally BS, Chaikledkaew U, Chaiyakunapruk N, Attia J, Thakkinstian A. Meta-analysis of economic evaluation studies: data harmonisation and methodological issues. *BMC Health Serv Res*. Feb 15, 2022;22(1):202. [[FREE Full text](#)] [doi: [10.1186/s12913-022-07595-1](https://doi.org/10.1186/s12913-022-07595-1)] [Medline: [35168619](https://pubmed.ncbi.nlm.nih.gov/35168619/)]
7. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. *Nature*. Mar 07, 2018;555(7695):175-182. [[FREE Full text](#)] [doi: [10.1038/nature25753](https://doi.org/10.1038/nature25753)] [Medline: [29517004](https://pubmed.ncbi.nlm.nih.gov/29517004/)]
8. Tsertsvadze A, Chen Y, Moher D, Sutcliffe P, McCarthy N. How to conduct systematic reviews more expeditiously? *Syst Rev*. Nov 12, 2015;4:160. [[FREE Full text](#)] [doi: [10.1186/s13643-015-0147-7](https://doi.org/10.1186/s13643-015-0147-7)] [Medline: [26563648](https://pubmed.ncbi.nlm.nih.gov/26563648/)]
9. Scott AM, Forbes C, Clark J, Carter M, Glasziou P, Munn Z. Systematic review automation tools improve efficiency but lack of knowledge impedes their adoption: a survey. *J Clin Epidemiol*. Oct 2021;138:80-94. [doi: [10.1016/j.jclinepi.2021.06.030](https://doi.org/10.1016/j.jclinepi.2021.06.030)] [Medline: [34242757](https://pubmed.ncbi.nlm.nih.gov/34242757/)]
10. Khalil H, Ameen D, Zarnegar A. Tools to support the automation of systematic reviews: a scoping review. *J Clin Epidemiol*. Apr 2022;144:22-42. [doi: [10.1016/j.jclinepi.2021.12.005](https://doi.org/10.1016/j.jclinepi.2021.12.005)] [Medline: [34896236](https://pubmed.ncbi.nlm.nih.gov/34896236/)]
11. Qureshi R, Shaughnessy D, Gill KAR, Robinson KA, Li T, Agai E. Are ChatGPT and large language models "the answer" to bringing us closer to systematic review automation? *Syst Rev*. Apr 29, 2023;12(1):72. [[FREE Full text](#)] [doi: [10.1186/s13643-023-02243-z](https://doi.org/10.1186/s13643-023-02243-z)] [Medline: [37120563](https://pubmed.ncbi.nlm.nih.gov/37120563/)]
12. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.4. Aug 2023. URL: <https://training.cochrane.org/handbook> [accessed 2024-06-11]
13. Deeks J, Bossuyt P, Leeflang M, Takwoingi Y. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*. Jul 2023. URL: <https://training.cochrane.org/handbook-diagnostic-test-accuracy/current> [accessed 2024-06-11]
14. Xiao Y, Watson M. Guidance on conducting a systematic literature review. *J Plan Educ Res*. Aug 28, 2017;39(1):93-112. [doi: [10.1177/0739456x17723971](https://doi.org/10.1177/0739456x17723971)]
15. Muka T, Glisic M, Milic J, Verhoog S, Bohlius J, Bramer W, et al. A 24-step guide on how to design, conduct, and successfully publish a systematic review and meta-analysis in medical research. *Eur J Epidemiol*. Jan 2020;35(1):49-60. [doi: [10.1007/s10654-019-00576-5](https://doi.org/10.1007/s10654-019-00576-5)] [Medline: [31720912](https://pubmed.ncbi.nlm.nih.gov/31720912/)]
16. Siddaway AP, Wood AM, Hedges LV. How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annu Rev Psychol*. Jan 04, 2019;70:747-770. [doi: [10.1146/annurev-psych-010418-102803](https://doi.org/10.1146/annurev-psych-010418-102803)] [Medline: [30089228](https://pubmed.ncbi.nlm.nih.gov/30089228/)]
17. Tawfik GM, Dila KAS, Mohamed MYF, Tam DNH, Kien ND, Ahmed AM, et al. A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Trop Med Health*. 2019;47:46. [[FREE Full text](#)] [doi: [10.1186/s41182-019-0165-6](https://doi.org/10.1186/s41182-019-0165-6)] [Medline: [31388330](https://pubmed.ncbi.nlm.nih.gov/31388330/)]
18. Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*. Feb 27, 2017;7(2):e012545. [[FREE Full text](#)] [doi: [10.1136/bmjopen-2016-012545](https://doi.org/10.1136/bmjopen-2016-012545)] [Medline: [28242767](https://pubmed.ncbi.nlm.nih.gov/28242767/)]

19. Michelson M, Reuter K. The significant cost of systematic reviews and meta-analyses: a call for greater involvement of machine learning to assess the promise of clinical trials. *Contemp Clin Trials Commun*. Dec 2019;16:100443. [FREE Full text] [doi: [10.1016/j.conctc.2019.100443](https://doi.org/10.1016/j.conctc.2019.100443)] [Medline: [31497675](https://pubmed.ncbi.nlm.nih.gov/31497675/)]
20. Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst Rev*. Jul 11, 2019;8(1):163. [FREE Full text] [doi: [10.1186/s13643-019-1074-9](https://doi.org/10.1186/s13643-019-1074-9)] [Medline: [31296265](https://pubmed.ncbi.nlm.nih.gov/31296265/)]
21. Clark J, Glasziou P, Del Mar C, Bannach-Brown A, Stehlik P, Scott AM. A full systematic review was completed in 2 weeks using automation tools: a case study. *J Clin Epidemiol*. May 2020;121:81-90. [doi: [10.1016/j.jclinepi.2020.01.008](https://doi.org/10.1016/j.jclinepi.2020.01.008)] [Medline: [32004673](https://pubmed.ncbi.nlm.nih.gov/32004673/)]
22. Luykx JJ, Gerritse F, Habets PC, Vinkers CH. The performance of ChatGPT in generating answers to clinical questions in psychiatry: a two-layer assessment. *World Psychiatry*. Oct 15, 2023;22(3):479-480. [FREE Full text] [doi: [10.1002/wps.21145](https://doi.org/10.1002/wps.21145)] [Medline: [37713576](https://pubmed.ncbi.nlm.nih.gov/37713576/)]
23. Roberts RH, Ali SR, Hutchings HA, Dobbs TD, Whitaker IS. Comparative study of ChatGPT and human evaluators on the assessment of medical literature according to recognised reporting standards. *BMJ Health Care Inform*. Oct 2023;30(1):e100830. [FREE Full text] [doi: [10.1136/bmjhci-2023-100830](https://doi.org/10.1136/bmjhci-2023-100830)] [Medline: [37827724](https://pubmed.ncbi.nlm.nih.gov/37827724/)]
24. Luo X, Estill J, Chen Y. The use of ChatGPT in medical research: do we need a reporting guideline? *Int J Surg*. Dec 01, 2023;109(12):3750-3751. [FREE Full text] [doi: [10.1097/JS9.0000000000000737](https://doi.org/10.1097/JS9.0000000000000737)] [Medline: [37707517](https://pubmed.ncbi.nlm.nih.gov/37707517/)]
25. Alshami A, Elsayed M, Ali E, Eltoukhy A, Zayed T. Harnessing the power of ChatGPT for automating systematic review process: methodology, case study, limitations, and future directions. *Systems*. Jul 09, 2023;11(7):351. [doi: [10.3390/systems11070351](https://doi.org/10.3390/systems11070351)]
26. Mahuli SA, Rai A, Mahuli AV, Kumar A. Application ChatGPT in conducting systematic reviews and meta-analyses. *Br Dent J*. Jul 2023;235(2):90-92. [doi: [10.1038/s41415-023-6132-y](https://doi.org/10.1038/s41415-023-6132-y)] [Medline: [37500847](https://pubmed.ncbi.nlm.nih.gov/37500847/)]
27. van Dijk SHB, Brusse-Keizer MGJ, Bucsán CC, van der Palen J, Doggen CJM, Lenferink A. Artificial intelligence in systematic reviews: promising when appropriately used. *BMJ Open*. Jul 07, 2023;13(7):e072254. [FREE Full text] [doi: [10.1136/bmjopen-2023-072254](https://doi.org/10.1136/bmjopen-2023-072254)] [Medline: [37419641](https://pubmed.ncbi.nlm.nih.gov/37419641/)]
28. Khraisha Q, Put S, Kappenberg J, Warratch A, Hadfield K. Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Res Synth Methods*. Mar 14, 2024:online ahead of print. [doi: [10.1002/jrsm.1715](https://doi.org/10.1002/jrsm.1715)] [Medline: [38484744](https://pubmed.ncbi.nlm.nih.gov/38484744/)]
29. Gwon YN, Kim JH, Chung HS, Jung EJ, Chun J, Lee S, et al. The use of generative AI for scientific literature searches for systematic reviews: ChatGPT and Microsoft Bing AI performance evaluation. *JMIR Med Inform*. May 14, 2024;12:e51187-e51187. [FREE Full text] [doi: [10.2196/51187](https://doi.org/10.2196/51187)] [Medline: [38771247](https://pubmed.ncbi.nlm.nih.gov/38771247/)]
30. Hossain MM. Using ChatGPT and other forms of generative AI in systematic reviews: challenges and opportunities. *J Med Imaging Radiat Sci*. Mar 2024;55(1):11-12. [doi: [10.1016/j.jmir.2023.11.005](https://doi.org/10.1016/j.jmir.2023.11.005)] [Medline: [38040497](https://pubmed.ncbi.nlm.nih.gov/38040497/)]
31. Giunti G, Doherty CP. Cocreating an automated mHealth apps systematic review process with generative AI: design science research approach. *JMIR Med Educ*. Feb 12, 2024;10:e48949. [FREE Full text] [doi: [10.2196/48949](https://doi.org/10.2196/48949)] [Medline: [38345839](https://pubmed.ncbi.nlm.nih.gov/38345839/)]
32. Nashwan AJ, Jaradat JH. Streamlining systematic reviews: harnessing large language models for quality assessment and risk-of-bias evaluation. *Cureus*. Aug 2023;15(8):e43023. [FREE Full text] [doi: [10.7759/cureus.43023](https://doi.org/10.7759/cureus.43023)] [Medline: [37674957](https://pubmed.ncbi.nlm.nih.gov/37674957/)]
33. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. *Am J Cancer Res*. 2023;13(4):1148-1154. [FREE Full text] [Medline: [37168339](https://pubmed.ncbi.nlm.nih.gov/37168339/)]
34. Issaiy M, Ghanaati H, Kolahi S, Shakiba M, Jalali AH, Zarei D, et al. Methodological insights into ChatGPT's screening performance in systematic reviews. *BMC Med Res Methodol*. Mar 27, 2024;24(1):78. [FREE Full text] [doi: [10.1186/s12874-024-02203-8](https://doi.org/10.1186/s12874-024-02203-8)] [Medline: [38539117](https://pubmed.ncbi.nlm.nih.gov/38539117/)]
35. Branum C, Schiavenato M. Can ChatGPT accurately answer a PICOT question? Assessing AI response to a clinical question. *Nurse Educ*. 2023;48(5):231-233. [doi: [10.1097/NNE.0000000000001436](https://doi.org/10.1097/NNE.0000000000001436)] [Medline: [37130197](https://pubmed.ncbi.nlm.nih.gov/37130197/)]
36. Macdonald C, Adeloye D, Sheikh A, Rudan I. Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis. *J Glob Health*. Feb 17, 2023;13:01003. [FREE Full text] [doi: [10.7189/jogh.13.01003](https://doi.org/10.7189/jogh.13.01003)] [Medline: [36798998](https://pubmed.ncbi.nlm.nih.gov/36798998/)]
37. Richard E, Pozzi A. Using ChatGPT to develop the statistical analysis plan for a randomized controlled trial: a case report. *Zurich Open Repository and Archive*. Preprint posted online on October 17, 2023. [doi: [10.21203/rs.3.rs-3433956/v1](https://doi.org/10.21203/rs.3.rs-3433956/v1)]
38. Hutson M. How AI is being used to accelerate clinical trials. *Nature*. Mar 2024;627(8003):S2-S5. [doi: [10.1038/d41586-024-00753-x](https://doi.org/10.1038/d41586-024-00753-x)] [Medline: [38480968](https://pubmed.ncbi.nlm.nih.gov/38480968/)]
39. Van der Mierden S, Tsaïoun K, Bleich A, Leenaars CHC. Software tools for literature screening in systematic reviews in biomedical research. *Altex*. 2019;36(3):508-517. [doi: [10.14573/altex.1902131](https://doi.org/10.14573/altex.1902131)] [Medline: [31113000](https://pubmed.ncbi.nlm.nih.gov/31113000/)]
40. Wang S, Scells H, Koopman B, Zucco G. Can ChatGPT write a good Boolean query for systematic review literature search? *arXiv*. Preprint posted online on February 9, 2023. [FREE Full text] [doi: [10.1145/3539618.3591703](https://doi.org/10.1145/3539618.3591703)]
41. Alaniz L, Vu C, Pfaff MJ. The utility of artificial intelligence for systematic reviews and Boolean query formulation and translation. *Plast Reconstr Surg Glob Open*. Oct 2023;11(10):e5339. [FREE Full text] [doi: [10.1097/GOX.0000000000005339](https://doi.org/10.1097/GOX.0000000000005339)] [Medline: [37908326](https://pubmed.ncbi.nlm.nih.gov/37908326/)]



42. Guimarães NS, Joviano-Santos JV, Reis MG, Chaves RRM, Observatory of Epidemiology, Nutrition, Health Research (OPENS). Development of search strategies for systematic reviews in health using ChatGPT: a critical analysis. *J Transl Med*. Jan 02, 2024;22(1):1. [FREE Full text] [doi: [10.1186/s12967-023-04371-5](https://doi.org/10.1186/s12967-023-04371-5)] [Medline: [38167166](https://pubmed.ncbi.nlm.nih.gov/38167166/)]
43. Tantry TP, Karanth H, Shetty PK, Kadam D. Self-learning software tools for data analysis in meta-analysis. *Korean J Anesthesiol*. Oct 2021;74(5):459-461. [FREE Full text] [doi: [10.4097/kja.21080](https://doi.org/10.4097/kja.21080)] [Medline: [33677944](https://pubmed.ncbi.nlm.nih.gov/33677944/)]
44. Cai X, Geng Y, Du Y, Westerman B, Wang D, Ma C, et al. Utilizing ChatGPT to select literature for meta-analysis shows workload reduction while maintaining a similar recall level as manual curation. medRxiv. Preprint posted online on September 7, 2023. [doi: [10.1101/2023.09.06.23295072](https://doi.org/10.1101/2023.09.06.23295072)]
45. Eugene S, Istvan D, Gauransh K. Assessing the ability of ChatGPT to screen articles for systematic reviews. arXiv. Preprint posted online on July 12, 2023. [doi: [10.5860/choice.189890](https://doi.org/10.5860/choice.189890)]
46. Kohandel Gargari O, Mahmoudi MH, Hajisafarali M, Samiee R. Enhancing title and abstract screening for systematic reviews with GPT-3.5 turbo. *BMJ Evid Based Med*. Jan 19, 2024;29(1):69-70. [FREE Full text] [doi: [10.1136/bmjebm-2023-112678](https://doi.org/10.1136/bmjebm-2023-112678)] [Medline: [37989538](https://pubmed.ncbi.nlm.nih.gov/37989538/)]
47. Guo E, Gupta M, Deng J, Park Y, Paget M, Naugler C. Automated paper screening for clinical reviews using large language models: data analysis study. *J Med Internet Res*. Jan 12, 2024;26:e48996. [FREE Full text] [doi: [10.2196/48996](https://doi.org/10.2196/48996)] [Medline: [38214966](https://pubmed.ncbi.nlm.nih.gov/38214966/)]
48. Polak MP, Morgan D. Extracting accurate materials data from research papers with conversational language models and prompt engineering. *Nat Commun*. Feb 21, 2024;15(1):1569. [doi: [10.1038/s41467-024-45914-8](https://doi.org/10.1038/s41467-024-45914-8)] [Medline: [38383556](https://pubmed.ncbi.nlm.nih.gov/38383556/)]
49. Mahmoudi H, Chang D, Lee H, Ghaffarzadegan N, Jalali MS. A critical assessment of large language models for systematic reviews: utilizing ChatGPT for complex data extraction. SSRN. Preprint posted online on April 19, 2024. [doi: [10.2139/ssrn.4797024](https://doi.org/10.2139/ssrn.4797024)]
50. Sun Z, Zhang R, Doi SA, Furuya-Kanamori L, Yu T, Lin L, et al. How good are large language models for automated data extraction from randomized trials? medRxiv. Preprint posted online on February 21, 2024. [doi: [10.1101/2024.02.20.24303083](https://doi.org/10.1101/2024.02.20.24303083)]
51. Marshall IJ, Kuiper J, Wallace BC. RobotReviewer: evaluation of a system for automatically assessing bias in clinical trials. *J Am Med Inform Assoc*. Jan 2016;23(1):193-201. [FREE Full text] [doi: [10.1093/jamia/ocv044](https://doi.org/10.1093/jamia/ocv044)] [Medline: [26104742](https://pubmed.ncbi.nlm.nih.gov/26104742/)]
52. Lai H, Ge L, Sun M, Pan B, Huang J, Hou L, et al. Assessing the risk of bias in randomized clinical trials with large language models. *JAMA Netw Open*. May 01, 2024;7(5):e2412687. [FREE Full text] [doi: [10.1001/jamanetworkopen.2024.12687](https://doi.org/10.1001/jamanetworkopen.2024.12687)] [Medline: [38776081](https://pubmed.ncbi.nlm.nih.gov/38776081/)]
53. Pitre T, Jassal T, Talukdar JR, Shahab M, Ling M, Zeraatkar D. ChatGPT for assessing risk of bias of randomized trials using the RoB 2.0 tool: a methods study. medRxiv. Preprint posted online on November 22, 2023. [doi: [10.1101/2023.11.19.23298727](https://doi.org/10.1101/2023.11.19.23298727)]
54. Rasheed Z, Waseem M, Ahmad A, Kemell KK, Xiaofeng W, Nguyen Duc A, et al. Can large language models serve as data analysts? A multi-agent assisted approach for qualitative data analysis. arXiv. Preprint posted online on February 2, 2024.
55. Kim S. Using ChatGPT for language editing in scientific articles. *Maxillofac Plast Reconstr Surg*. Mar 08, 2023;45(1):13. [FREE Full text] [doi: [10.1186/s40902-023-00381-x](https://doi.org/10.1186/s40902-023-00381-x)] [Medline: [36882591](https://pubmed.ncbi.nlm.nih.gov/36882591/)]
56. Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digit Med*. Apr 26, 2023;6(1):75. [doi: [10.1038/s41746-023-00819-6](https://doi.org/10.1038/s41746-023-00819-6)] [Medline: [37100871](https://pubmed.ncbi.nlm.nih.gov/37100871/)]
57. Kim JK, Chua M, Rickard M, Lorenzo A. ChatGPT and large language model (LLM) chatbots: the current state of acceptability and a proposal for guidelines on utilization in academic medicine. *J Pediatr Urol*. Oct 2023;19(5):598-604. [doi: [10.1016/j.jpuro.2023.05.018](https://doi.org/10.1016/j.jpuro.2023.05.018)] [Medline: [37328321](https://pubmed.ncbi.nlm.nih.gov/37328321/)]
58. Mugaanyi J, Cai L, Cheng S, Lu C, Huang J. Evaluation of large language model performance and reliability for citations and references in scholarly writing: cross-disciplinary study. *J Med Internet Res*. Apr 05, 2024;26:e52935. [FREE Full text] [doi: [10.2196/52935](https://doi.org/10.2196/52935)] [Medline: [38578685](https://pubmed.ncbi.nlm.nih.gov/38578685/)]
59. Firdausi Nuzula I, Miftahul Amri M. Will ChatGPT bring a new paradigm to HR World? A critical opinion article. *J Manag Stud Devel*. Apr 24, 2023;2(02):142-161. [doi: [10.56741/jmsd.v2i02.316](https://doi.org/10.56741/jmsd.v2i02.316)]
60. Eriksen MB, Frandsen TF. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *J Med Libr Assoc*. Oct 2018;106(4):420-431. [FREE Full text] [doi: [10.5195/jmla.2018.345](https://doi.org/10.5195/jmla.2018.345)] [Medline: [30271283](https://pubmed.ncbi.nlm.nih.gov/30271283/)]
61. Roberts R. I tested how well ChatGPT can pull data out of messy PDFs. Source. Mar 01, 2023. URL: <https://source.opennews.org/articles/testing-pdf-data-extraction-chatgpt/> [accessed 2024-06-11]
62. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, et al. Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. Oct 18, 2011;343:d5928. [doi: [10.1136/bmj.d5928](https://doi.org/10.1136/bmj.d5928)] [Medline: [22008217](https://pubmed.ncbi.nlm.nih.gov/22008217/)]
63. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. Aug 28, 2019;366:l4898. [doi: [10.1136/bmj.l4898](https://doi.org/10.1136/bmj.l4898)] [Medline: [31462531](https://pubmed.ncbi.nlm.nih.gov/31462531/)]

64. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. Mar 29, 2021;372:n71. [[FREE Full text](#)] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
65. Gaggioli A. Ethics: disclose use of AI in scientific manuscripts. *Nature*. Feb 14, 2023;614(7948):413-413. [doi: [10.1038/d41586-023-00381-x](https://doi.org/10.1038/d41586-023-00381-x)] [Medline: [36788370](https://pubmed.ncbi.nlm.nih.gov/36788370/)]
66. Wang L, Ma C, Feng X, Zhang Z, Yang H, Zhang J, et al. A survey on large language model based autonomous agents. *Front Comput Sci*. Mar 22, 2024;18(6):186345. [[FREE Full text](#)] [doi: [10.1007/s11704-024-40231-1](https://doi.org/10.1007/s11704-024-40231-1)]
67. Hong S, Zhuge M, Chen J, Zheng X, Cheng Y, Zhang C, et al. Metagpt: Meta programming for multi-agent collaborative framework. *arXiv*. Preprint posted online on November 6, 2023.
68. Flanagin A, Pirracchio R, Khera R, Berkwits M, Hswen Y, Bibbins-Domingo K. Reporting use of AI in research and scholarly publication-JAMA Network Guidance. *JAMA*. Apr 02, 2024;331(13):1096-1098. [doi: [10.1001/jama.2024.3471](https://doi.org/10.1001/jama.2024.3471)] [Medline: [38451540](https://pubmed.ncbi.nlm.nih.gov/38451540/)]

## Abbreviations

**AI:** artificial intelligence

**API:** application programming interface

**LLM:** large language model

**PICO:** Population, Intervention, Comparison, Outcome

**PRISMA:** Preferred Reporting Items for Systematic reviews and Meta-Analyses

**RoB:** Risk of Bias

*Edited by G Eysenbach; submitted 20.02.24; peer-reviewed by A Jafarizadeh, M Chatzimina, AS Van Epps; comments to author 03.05.24; revised version received 21.05.24; accepted 29.05.24; published 25.06.24*

*Please cite as:*

*Luo X, Chen F, Zhu D, Wang L, Wang Z, Liu H, Lyu M, Wang Y, Wang Q, Chen Y*

*Potential Roles of Large Language Models in the Production of Systematic Reviews and Meta-Analyses*

*J Med Internet Res 2024;26:e56780*

*URL: <https://www.jmir.org/2024/1/e56780>*

*doi: [10.2196/56780](https://doi.org/10.2196/56780)*

*PMID: [38819655](https://pubmed.ncbi.nlm.nih.gov/38819655/)*

©Xufei Luo, Fengxian Chen, Di Zhu, Ling Wang, Zijun Wang, Hui Liu, Meng Lyu, Ye Wang, Qi Wang, Yaolong Chen. Originally published in the *Journal of Medical Internet Research* (<https://www.jmir.org>), 25.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet Research* (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.